and were seated in a primate chair with their head fixed, facing a computer monitor that presented images on each trial. The researchers isolated and monitored single neurons in the lateral prefrontal cortex (LPFC), an area of the brain associated with intentional, purposive action (operant behavior), and reinforced neuronal spikes that exceeded an established criterion. When the neuron emitted a spike in the established range, the monkeys received juice in the mouth as reinforcement. Control conditions involved all of the procedures of the experimental phase, including visual fixation and presentation of juice, but there was no contingency between neuron firing and reinforcement. Evidence indicated that individual neurons in the LPFC showed operant conditioning. Further experiments in the series indicated that LPFC neurons would respond in accord with the schedule of reinforcement. The researchers noted that alternative explanations such as simple reward prediction, attention, and arousal were unlikely to account for the findings. Also, the LPFC has few motor neurons, so it is unlikely that neuron activity in this region directly coincides with movement of the animal. One possibility is that LPFC neurons contribute to behavioral flexibility—a neural substrate for operant behavior. These neurons may "enhance the action signals [in other brain areas] to ensure the motor execution when operant control is required" (Kobayashi et al., 2010, p. 1854; see also Schafer & Moore, 2011; see also Ishikawa, Matsumoto, Sakaguchi, Matsuki, & Ikegaya, 2014 for "*in vivo*" rapid operant conditioning of selected hippocampal neurons in mice, using neurofeedback reinforcement that requires NMDA receptor activity—extending the "*in-vitro*" experiments described previously).

Operant conditioning is a major adaptive mechanism for animals that change behavior on the basis of lifetime experiences (ontogeny). From a biological standpoint, operant conditioning allows for behavioral flexibility, survival, and reproduction. Evidence is accumulating that behavioral flexibility is based on **neuroplasticity**—alterations of neurons and neural interconnections during a lifetime by changes in the environmental contingencies (Caroni, Donato, & Muller, 2012). *In-vitro* reinforcement experiments show that endogenous brain chemicals binding to particular types of receptors increase the likelihood of neuronal activity. These molecular neural processes presumably underlie the large-scale changes in operant behavior that occur as humans and other animals interact with the world in which they live, from moment to moment, over a life span.

Procedures in Operant Conditioning

Operant Rate and Probability of Response

Rate of response refers to the number of operant responses that occur in some defined unit of time. For example, if you ask 5 questions during a 2-h class, your rate is 2.5 questions per hour. An animal that presses a lever 1000 times in a 1-h session generates a rate of 1000 bar presses per hour (or 16.7 responses per minute). Skinner (1938) proposed that rate of response is the basic datum (or measure) for operant analysis. **Operant rate** is a measure of the probability of behavior (the **probability of response**). In other words, an operant that occurs at a high rate in one situation has a high probability of being emitted in a similar situation in the future. This increased probability of response is observed as a change in operant rate. Of course, probability of response may decrease, and in this case is observed as a reduction in rate.

The Free-Operant Method

> The Skinner box is not a specific technique, but rather a method of research employing the free operant. The use of the free operant is a method of wide generality; it refers to any apparatus that generates a response which takes a short time to occur and leaves the animal in the same place ready to respond again. The free operant is used in experiments when the main dependent variable is the frequency of occurrence of behavior. Nearly all the problems of a science of behavior fit this paradigm when the questions are of the form: what is the likelihood of a piece of behavior occurring on this particular occasion; how strong is the tendency to behave on this occasion relative to another occasion? The free operant has advantages in this respect, because it removes restrictions on the frequency with which a response can occur and permits the observation of moment-to-moment changes in frequency.
> (Ferster, 1953, p. 263)

In the **free-operant method**, an animal may repeatedly respond over an extensive period of time. The organism is free to emit many responses or none at all. Using this method, responses can be made without interference from the experimenter. For example, a laboratory rat may press a lever for food pellets. Lever pressing is under the control of the animal, which may press the bar rapidly, slowly, or quit pressing. Importantly, this method allows the researcher to observe changes in the *rate of response*. This is important because rate of response is used as a measure of response probability. Rate of response must be free to vary if it is used to index the future probability of operant behavior.

The analysis of operant rate and probability of response is not easily accomplished when an organism is given a series of trials (as in the Thorndike experiments). This is because the experimenter largely controls the animal's rate of behavior. For example, a rat that runs down a T-maze for food reward is picked up at the goal box and returned to the starting point. Because the experimenter sets the number of trials and response opportunities, changes in the rate of response cannot be directly observed and measured. Comparing the T-maze trial procedure with the free-operant method, it is clear that the free-operant method is more suitable for studying the *probability of response* in a given situation. The free-operant method is clearly demonstrated by the procedures used in operant conditioning.

The Operant Chamber

To study operant conditioning in a laboratory, a device called an **operant chamber** is used. Of course, operant conditioning is also investigated outside laboratories. Nonetheless, investigating the behavior of animals in operant chambers has resulted in the discovery of many principles of behavior. Figure 4.7 shows a student setup of an operant chamber designed to accommodate a laboratory rat (note that a research setup would involve much more experimental control, such as a sound-attenuating enclosure and "white noise" to mask sounds from outside). The chamber is a small enclosed box that contains a lever with a light above it, and a food magazine or cup connected to an external feeder. The feeder delivers a small food pellet (typically 45 mg) when electronically activated. In this situation, the food pellet serves as reinforcement for lever pressing. The operant chamber *structures the situation* so that the desired behavior will occur and incompatible behavior is reduced. Thus, lever pressing is highly likely, while behavior such as running away is minimized. A school classroom also attempts to structure the behavior of students with regard to learning. The classroom, unlike the operant chamber, often contains many distractions (e.g., looking out the window) that interfere with on-task behavior and concentrating on the material being presented.

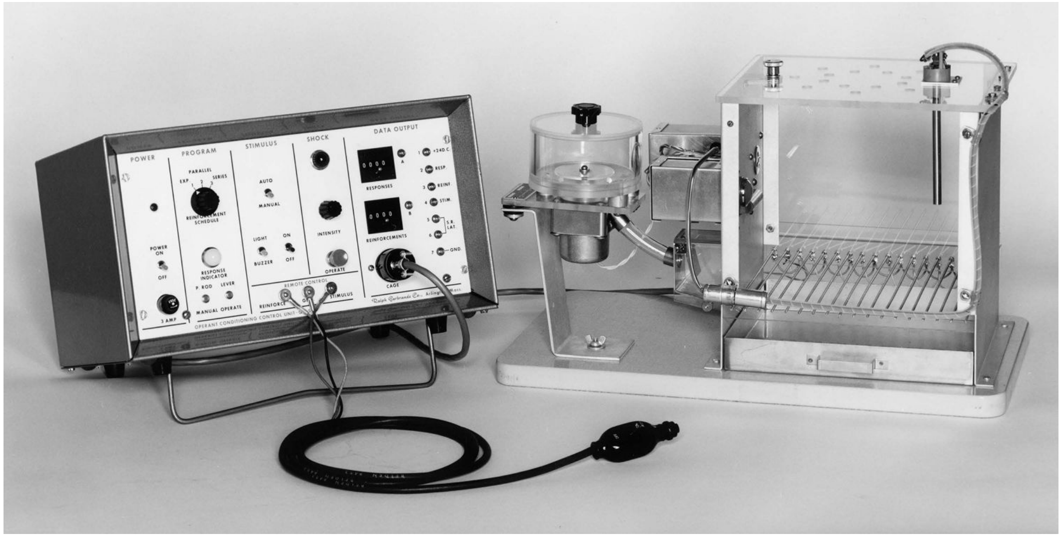

FIG. 4.7 A student setup of an operant chamber for a rat. The chamber is a small box that has a lever that the animal can press. There is a light above the lever that can be turned on or off. A food magazine or cup is connected to an electronically activated feeder. The feeder delivers a small, 45-mg food pellet to the cup. In this situation, the food pellet serves as reinforcement for lever pressing. *Source*: Ralph Gerbrands Corporation, Arlington, MA.

Establishing Motivation: Deprivation

Because the delivery of food is used as reinforcement, an animal must be motivated to obtain food. An objective and quantifiable measure of motivation for food is percentage of free-feeding body weight (note that another way of quantifying deprivation is a timed cycle that specifies the time since the rat last consumed the reinforcer). Prior to a typical experiment, an animal is brought from a commercial (or research) colony into a laboratory, placed in a cage, given free access to food, and weighed on a daily basis. The average weight is calculated, and this value is used as a baseline. Next, the daily food ration is reduced until the animal reaches 85% of its free-feeding weight. The procedure of restricting access to food (the potentially reinforcing stimulus) is called a **deprivation operation** (see "Motivational Operations" in Chapter 2). At this point, the experimenter assumes, but does not know, that food is a reinforcing stimulus. This is because food delivery must increase the frequency of an operant before it can be defined as reinforcement.

The weight loss or deprivation criterion is less severe than it first appears. Laboratory animals typically have food freely available 24 h a day, whereas animals in the wild must forage for their food. The result is that lab animals tend to be heavier than their free-ranging counterparts. Alan Poling and his colleagues nicely demonstrated this point by showing that captured free-range pigeons gained an average 17% body weight when housed under laboratory conditions (Poling, Nickel, & Alling, 1990). Notice that weight gain, for these birds, is roughly equal to the weight loss typically imposed on laboratory animals.

At the physiological level, loss of body weight (food deprivation) activates hormones related to energy homeostasis, involving insulin, leptin, and ghrelin. Neuroendocrine research shows that variation in these hormones increase (or decrease) food-reinforced behavior and modulate neurotransmitters controlling the endogenous reward circuitry, especially the midbrain dopamine (DA) and opiodergic systems. Thus, changes in the feeding regime of the rat impact operant control by reinforcement through a complex interplay of metabolism, neurotransmitters, and endogenous reward pathways (Figlewicz & Sipols, 2010).

Magazine Training

After food deprivation is established, **magazine training** starts. For example, a rat is placed in an operant chamber and a microcomputer periodically turns on the feeder. When the feeder is turned on, it makes a click and a small 45-mg food pellet falls into the food magazine. Because the click and the appearance of food are associated in time, you would, after training, observe a typical rat staying close to the food magazine; also, the animal would move quickly toward the magazine when the feeder operated and the click occurred. Because the click of the feeder reliably precedes the appearance of food, it becomes a conditioned positive reinforcer (see Chapter 10 on conditioned reinforcement). A **conditioned reinforcer** is an event or stimulus that acquires its reinforcing function over the lifetime of the organism (ontogeny). In this case, following the click of the feeder by the presentation of food establishes the sound of the feeder as a conditioned reinforcer for the rat.

The Operant Class

Staying close to the food cup and moving toward it are operants that have been selected by their reinforcing consequences. In other words, these responses have been reliably followed by food presentation, and as a result they increased in frequency. But, hovering around a food cup and moving toward it are operants that are difficult to measure objectively. In contrast, a lever press may be easily defined as a switch closure that makes an electrical connection. Any behavior emitted by the rat that results in a switch closure defines the operant class. A lever press with the left or right paw produces an identical electrical connection. Another advantage of lever pressing as an operant is that it may be emitted at high or low rates of response. This is an advantage because the primary focus of operant research is on the conditions that affect the rate (probability) of operant behavior.

Operant Level and Continuous Reinforcement

After magazine training, the food-deprived rat is again placed in the operant chamber. The researcher may first want to measure the rate of lever pressing before these responses produce food pellets. Rats emit many exploratory and manipulative responses, and as a result may press the lever at some low frequency, even when this behavior is not reinforced with food. This baseline rate of response is called the **operant level** or the rate of response before any known conditioning. Next, the environment is arranged so that each lever press results in the click of the feeder (conditioned reinforcement) and the delivery of a food pellet (primary reinforcement). When each response produces food, the schedule of reinforcement is called **continuous reinforcement (CRF)**. The food pellets are contingent on lever pressing. This contingency between the operant behavior and food reinforcement increases the frequency of lever pressing above operant level.

Shaping: The Method of Successive Approximation

In the preceding example, we took advantage of a rat's behavioral repertoire. The animal's **repertoire** refers to the behavior it is capable of naturally emitting on the basis of species and environmental history. Suppose you want to train some response that the animal does not emit. For example, you may want the rat to activate a switch by an upward thrust of its nose. A baseline period of observation shows that the animal fails to emit this response (in other words, the *operant level is zero*). In this case, the researcher could use **shaping** or the method of **successive approximation** to establish the response. This method involves reinforcing closer and closer approximations to the final performance (i.e., nosing the lever).

Steps in shaping lever pressing

- Standing near the lever reinforced
- Facing lever reinforced
- Move head toward lever reinforced
- Touching lever with parts of body (nose) reinforced
- Touching lever with paws reinforced
- Touching lever with specified (right) paw reinforced
- Partially depressing lever with right paw reinforced
- Raising and depressing lever with right paw reinforced
- Raising and depressing lever completely with right paw reinforced

FIG. 4.8 Figure depicts the possible steps in shaping the rat toward the final performance of pressing the lever with sufficient force to activate the feeder and deliver the food pellet.

We begin to shape lever pressing for food (final performance) with a rat that has already been conditioned to eat from the pellet feeder. At first, the rat is reinforced for standing in the vicinity of the lever. It is important to note that the most immediate consequence is the sound of the pellet feeder and it is this *conditioned reinforcer* which is used to shape the desired response.

Once the rat is reliably facing the lever, a movement of the head toward the lever is reinforced with a click of the feeder and presentation of food. Next, closer and closer approximations to depressing the lever with the right paw are reinforced, as shown in Figure 4.8. Each step of the procedure involves reinforcing closer approximations to the final performance and no reinforcement (extinction) of more distant responses. Eventually, the rat consistently emits responses with sufficient force to activate the electrical switch (*final performance*).

Shaping of key pecking in the pigeon uses the same principles as shaping of lever pressing in a rat. Skinner (1953) described how to train a bird to peck a spot in its cage using successive approximation to the final performance. Skinner emphasizes how the distribution of responses for the response class of “pecking a spot for food” is moved from an initial near zero probability toward a high and stable rate of response:

> We first give the bird food when it turns slightly in the direction of the spot from any part of the cage. This increases the frequency of such behavior. We then withhold reinforcement until a slight movement is made toward the spot. This again alters the general distribution of behavior without producing a new unit. We continue by reinforcing positions successively closer to the spot, then by reinforcing only when the head is moved slightly forward, and finally only when the beak actually makes contact with the spot. . . . The original probability of the response in its final form is very low; in some cases, it may even be zero. In this way we can build complicated operants, which would never appear in the repertoire of the organism otherwise. By reinforcing a series of successive approximations, we bring a rare response to a very high probability in a short time. . . . The total act of turning toward the spot from any point in the box, walking toward it, raising the head, and striking the spot may seem to be a functionally coherent unit of behavior; but it is constructed by a continual process of differential reinforcement from undifferentiated behavior, just as the sculptor shapes his figure from a lump of clay.
>
> (Skinner, 1953, pp. 92–93)

Notice that shaping makes use of **behavioral variability**—the animal’s tendency to emit variations in response form in a given situation. The range of behavioral variation is related to an animal’s capabilities based on genetic endowment, degree of neuroplasticity, and previous interactions with the environment. Behavioral variability allows for selection by reinforcing consequences, and is analogous to the role of genetic variability in natural selection (Neuringer, 2009, p. 319). Shaping by successive approximation uses undifferentiated operant behavior (a lump of clay), which is moved toward “a functionally coherent unit” (pecking a spot for food) by a process of **ontogenetic selection** (differential reinforcement). Shaping as an important behavioral procedure cannot be overemphasized. It is the process by which nearly all complex behaviors are acquired. The steps involve the explicit definition of the terminal behavior (final performance), description of the subject's or student's current repertoire (baseline level), and the contingent delivery of an effective reinforcer after each approximation toward the terminal performance (see also "Behavior Analysis in Education," Chapter 13).

NEW DIRECTIONS: Shaping and Neuroscience of Birdsong

Successive approximation to a final performance occurs as part of the natural contingencies shaping the singing of songbirds. The learning of birdsong appropriate to one's species begins with auditory-vocal correspondence; the nestling hears the song of adult birds (tutors), which evokes species-specific, highly variable vocal responses or subsong, akin to babbling in human infants (Brainard & Doupe, 2002). Research indicates that this "sensory learning" phase appears to involve activation and encoding by the mirror-neuron system (Prather, Peters, Nowicki, & Mooney, 2008; see also Chapter 11 on correspondence and mirror neurons).

To achieve adult song, nestlings also must be able to hear themselves. Birds deafened after exposure to the songs of tutors, but before they practice singing on their own, show abnormal songs as adults. The songs of the nestlings must be perfected by self-initiated practice (operant) and auditory feedback (reinforcement)—called the "sensorimotor learning" phase. In this phase, immature birdsong is shaped toward the adult song (final performance) by hearing how closely vocal responses correspond with those of adult birds (auditory feedback). Using self-produced feedback from singing, the youngster's melody is fine-tuned toward an adult "crystallized" song with only small variations from one rendition to the next.

At the neural level, vocal learning in birds is viewed as a form of motor learning, involving the *anterior forebrain pathway* (AFP) composed of an area homologous to the mammalian basal ganglia, as well as an area of the thalamus. Basal ganglia pathways connect with the *posterior descending pathway* (PDP), especially the high vocal center (HVC), to regulate production, timing, and sequencing of song. To achieve mature birdsong, neural circuitry is required for production of vocalizations, circuitry for hearing and discriminating sounds of self and others, and connections between the two pathways (Brainard & Doupe, 2013). Some researchers propose an "error correction" model for acquisition (and maintenance) of birdsong, emphasizing the AFP and the lateral magnocellular nucleus of anterior nidopallium (LMAN). In this model, the LMAN, by its connection to the premotor nucleus (RA), allows for adjustments in song by hearing self-produced vocalizations (perhaps encoded by the nidopallium caudomedial or NCM, an analog of the mammalian audio-association cortex) compared with the "song template" of the adult birds. At the present time, however, the exact neural processes involved in the use of auditory feedback to shape mature birdsong are not well understood (Brainard & Doupe, 2013). An operant delayed matching to sample analysis (see Chapter 8), with adult song as the sample and hearing self-produced song as the matching response (via LMAN, NCM, and RA circuitry), would emphasize the "hits" (reinforcement) for selection of neural connections and pathways (plasticity) rather than appealing to an inferred cognitive process, involving comparison of errors between self-song and the memory song template (see "Focus On: Behavioral Neuroscience and Operant Conditioning of the Neuron" in this chapter).

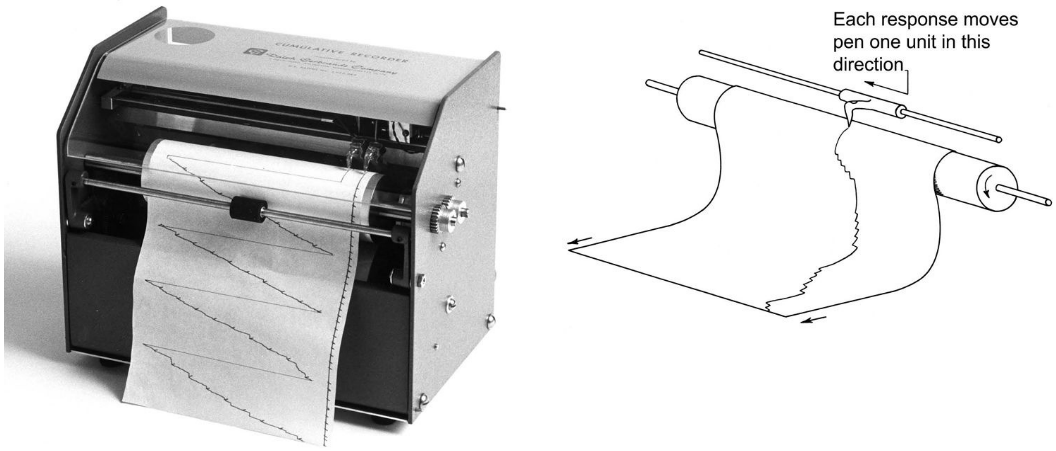

FIG. 4.9 Photograph shows a laboratory instrument used to record operant responses, called a cumulative recorder. This device was used in the early research on operant conditioning. The recorder gives a real-time measure of the rate of operant behavior. The faster the lever presses, the steeper the slope or rise of the cumulative record. This occurs because paper is drawn across the roller at a constant speed and the pen steps up a defined distance for each response. Modern laboratories show response records on computer screens along with additional behavioral measures.
Source: Ralph Gerbrands Corporation, Arlington, MA.

Recording Operant Behavior

A commonly used laboratory instrument that records the frequency of operant behavior in time (rate) is called a **cumulative recorder**. Figure 4.9 illustrates this device, which was used in the early research on operant conditioning. Each time a lever press occurs, the pen steps up one increment. When reinforcement occurs, this same pen makes a downward deflection. Once the pen reaches the top of the paper, it resets to the bottom and starts to step up again. Since the paper is drawn across the roller at a constant speed, the cumulative recorder depicts *a real-time measure of the rate of operant behavior*. The faster the lever presses, the steeper the slope or rise of the **cumulative record**.

A cumulative record of key pecking by a pigeon is shown in Figure 4.10. In this illustration, a bird responded 50 times to produce one food delivery. Notice that periods of responding are followed by reinforcement (indicated by the deflection of the pen). After reinforcement, the rate of response is zero, as indicated by the plateaus or flat portions of the cumulative record.

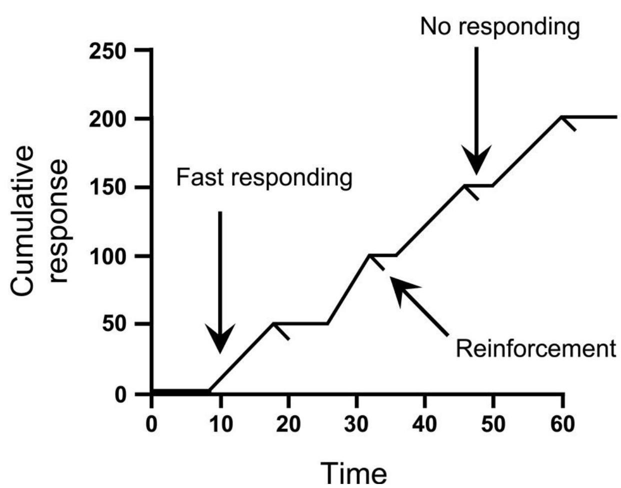

FIG. 4.10 A cumulative record of key pecking by a pigeon. In this illustration, a bird responded 50 times to produce one food delivery. Notice that 50 pecks are followed by reinforcement and that this is indicated by a downward deflection of the pen. Following reinforcement, the rate of response is zero, as indicated by the plateaus or flat portions of the record.

In a modern operant laboratory, the cumulative record is portrayed on a computer screen to provide the experimenter with an immediate report of the animal's behavior. Researchers have discovered many basic principles of behavior by examining cumulative records (e.g., Ferster & Skinner, 1957). Today, microcomputers allow researchers to collect, display, and record multiple measures of behavior (e.g., rate of response) that are later submitted to complex numerical analyses (Chen & Steinmetz, 1998; Gollub, 1991). In this book, we present examples of cumulative records and numerical analyses that have been important to the experimental analysis of behavior.

A Model Experiment

In the previous discussion of operant behavior, some basic principles were illustrated using the laboratory rat. It is important to realize that these same principles can be extended to a variety of species, including humans. In the following demonstration of operant conditioning, pigeons are used as the experimental subjects. Pigeons are placed in an operant chamber and required to peck a small plastic disk or *key* illuminated by a white light. A peck at the key activates a microswitch and makes an electrical connection that controls a food hopper. Presentation of food functions as reinforcement for pecking. A food hopper filled with grain swings forward and remains available for a few seconds. The bird can eat the grain by sticking its head through an opening. Figure 4.11 shows an operant chamber designed for birds. Note that the chamber is very similar to the one used to study the operant behavior of rats.

Before an experiment, the bird is taken from its home colony and placed alone in a cage. Each pigeon is given free access to food and water. The bird is weighed each day for about a week, and its baseline weight is calculated. Next, the daily food ration is reduced until the bird reaches approximately 80% of free-feeding or **ad libitum weight**. After the deprivation procedure, the pigeon is placed in the operant chamber for magazine training.

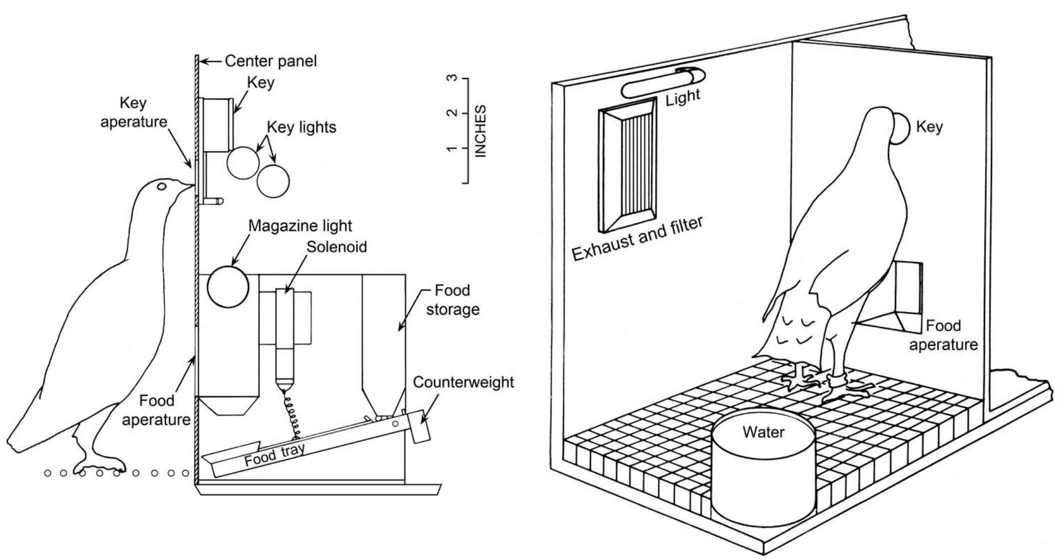

FIG. 4.11 An operant chamber for birds. The chamber contains a small plastic disk illuminated by a light. A peck at the disk activates a microswitch and makes an electrical connection. When reinforcement is scheduled to occur, the food hopper swings forward and remains available for a few seconds. The bird can eat grain from the hopper by sticking its head through the opening in the chamber wall. In principle, the chamber is similar to the one used to study the operant behavior of rats. *Source*: Adapted from C. B. Ferster & B. F. Skinner (1957), *Schedules of reinforcement*. New York: Appleton-Century-Crofts.

When the bird is placed in the chamber for the first time, it may show a variety of emotional responses, including wing flapping and defecating. This is because the chamber presents a number of novel features that initially function as aversive stimuli. For example, the operation of the feeder makes a loud sound that may startle the bird. Eventually, these emotional responses are extinguished by repeated exposure to the apparatus. As the emotional responses dissipate, the bird explores the environment and begins to eat from the food magazine. Since the sound of the hopper is associated with food, the sound becomes a conditioned positive reinforcer. At this point, the bird is said to be magazine trained.

The purpose of this demonstration is to train the pigeon to peck the key for food reinforcement. To show that the behavior occurs because of the contingency between pecking and food, an operant level or baseline of pecking the key must be measured. This is accomplished by placing the bird in the operant chamber and recording pecks on the key before a peck–food contingency is established. In other words, pecking the key does not produce food during the baseline phase of the experiment. The *operant level* serves as a baseline or control period for assessing a change in behavior.

A bird's operant level of key pecking is typically very low, and it is convenient to train these responses by the method of successive approximation. Shaping of key pecking in pigeons is similar to shaping lever pressing in rats; in both cases, shaping involves reinforcing closer and closer approximations to the final performance of pecking the key hard enough to operate the microswitch. As each approximation occurs, it is reinforced with the presentation of the food hopper. Earlier approximations are no longer reinforced and reduce in frequency (extinction). This process of reinforcing closer approximations, and withholding reinforcement for earlier approximations, eventually results in the pigeon pecking the key with sufficient force to operate the microswitch. The key peck that operates the microswitch to produce food is the *first definable response*. The switch closure and electrical connection define the operant class of pecking for food. At this point, a microcomputer is programmed so that each key peck results in the presentation of food for a few seconds. Because each response produces reinforcement, the schedule is called continuous reinforcement, or CRF.

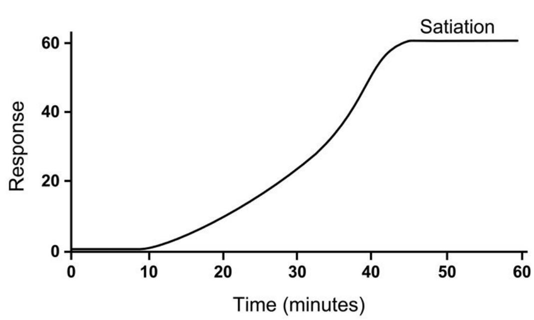

FIG. 4.12 Typical acquisition of key pecking on CRF or continuous reinforcement. Because every response is reinforced, downward deflections indicating reinforcement are omitted. Rate of response is low when the animal is initially placed in the chamber. After this brief period, rate of response is high and stable. Finally, rate of response declines and then levels off. This latter effect is caused by satiation.

Figure 4.12 shows the acquisition of key pecking on continuous reinforcement (the bird has presumably been shaped to peck the key for food). Notice that the rate of response is low when the pigeon is initially placed in the chamber. This period is called the warm-up, and probably occurs because of the abrupt change from home cage to the operant chamber. After the brief warm-up period, the rate of response is high and stable.

Finally, the record shows that rate of response declines, and the plateau indicates that the bird stops pecking the key. This latter effect is called **satiation**, and it occurs because the bird has eaten enough food. More technically, *rate of response declines because repeated presentations of the reinforcer weaken its effectiveness*. A satiation operation decreases the effectiveness of reinforcement. This effect is opposite to deprivation in which withholding the reinforcer increases its effectiveness.

To be sure that an increase in the rate of response is caused by the contingency of reinforcement, it is necessary to withdraw that contingency. In other words, if food is no longer presented, the pigeon should give up pecking the key. If the peck–food contingency caused key pecking, then withdrawal of the contingency will result in a decline in key pecking toward the operant level.

Figure 4.13 presents cumulative records for periods in which pecking produces or does not produce food. The initial peck–food contingency produces a steady rate of response. When pecking no longer produces food, the rate of response declines and eventually key pecking stops. Thus, key pecking clearly depends upon the contingency of reinforcement.

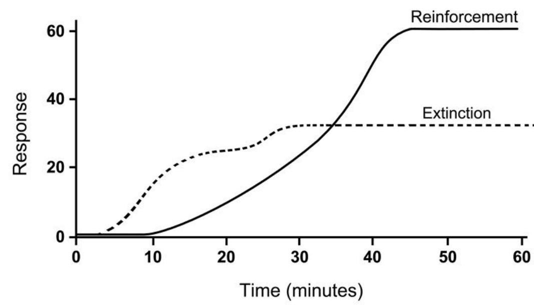

FIG. 4.13 Performances on CRF and extinction are shown. Responses are maintained when reinforced. When responding is no longer reinforced, however, rate of response declines and eventually responding stops.

FOCUS ON: Reinforcement and Problem Solving

Reinforcement and Response Stereotypy

Barry Schwartz (1980, 1982a) carried out a series of experiments with pigeons to show that reinforcement produced *response stereotypy*, a pattern of responding that occurred over and over again. Subsequently, Schwartz (1982b) used similar procedures with college students to demonstrate the presumed negative effects of reinforcement for human problem solving.

College students were given points on a counter when they completed a complex sequence of responses. The responses were left and right key presses that moved a light on a checkerboard-like matrix of 25 illuminated squares. Figure 4.14 shows the matrix, with the light in the top left square. The task required that the participant press the keys to move the light from the top left corner to the bottom right square. A press on the right key moved the light one square to the right. When the left-hand key was pressed, the light moved one square down. Schwartz required exactly four left (L) and four right (R) presses in any order (e.g., LRLRLRLR, LLLLRRRR, etc.). There were 72 different orders of left and right key presses that would move the light to the bottom right corner. When the light reached the bottom right corner, a point registered on the counter. The points were later exchanged for money. If the participant pressed any key a fifth time (e.g., RRRRR), all of the matrix lights were turned off and the trial ended without reinforcement.

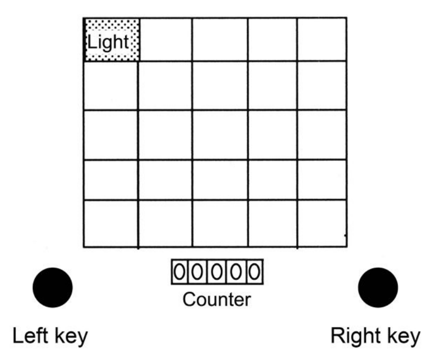

FIG. 4.14 The matrix task used by Schwartz (1982b). A right press moved the light one square to the right; a left button press moved the light down one square.

In a series of experiments, Schwartz found that students developed a stereotyped pattern of responding. The point is that as soon as a student hit on a correct sequence, he or she repeated it and rarely tried another pattern. In other experiments (Schwartz, 1982b), participants were explicitly reinforced for varying their response pattern. When this was done, the students developed higher-order stereotypes. From these experiments, Schwartz concluded that reinforcement interfered with problem solving because it produced stereotyped response patterns.

Reinforcement and Response Variability

Allen Neuringer (Figure 4.15) is a behavior analyst who investigates variability, randomness, and behavior (see Neuringer & Jensen, 2013 for an overview of this research). He suggested that the contingencies of reinforcement of the Schwartz experiments produced response stereotypy, but this was not an inevitable outcome of reinforcement. In the Schwartz experiments, the response patterns were constrained by the requirement to emit exactly four pecks on each key in any order. This constraint means that of the 256 possible sequences of pecks, only 72 patterns resulted in reinforcement; also, a fifth peck to either left or right keys resulted in a *timeout from reinforcement* (negative punishment). This timeout contingency punished many instances of response variability, reducing its occurrence.

A classic study by Page and Neuringer (1985) eliminated the constraints imposed by Schwartz and tested the assumption that variability could actually increase with reinforcement. The experiment involved pigeons pecking left and right keys on the light matrix task. Each session consisted of numerous trials with the bird emitting 8 pecks to the two keys on each trial—ending with food for 8 pecks that met the criterion for variability and a short timeout for other sequences. The contingency for the variability phase (VAR) involved lags in which the current sequence of 8 pecks had to differ from the pattern of pecks emitted on some previous trials. The number of previous trials defined the Lag value. For example, using a Lag 3 contingency, the current sequence of 8 pecks to left and right keys had to diverge from each of the patterns for the previous three trials for the current trial to end in food reinforcement; other "incorrect" sequences of pecks resulted in a timeout. Birds met the schedule requirements for variability even at Lag 50 by responding almost randomly (like coin flipping) on the left and right keys, not by "remembering" what they did over the past 50 trials.

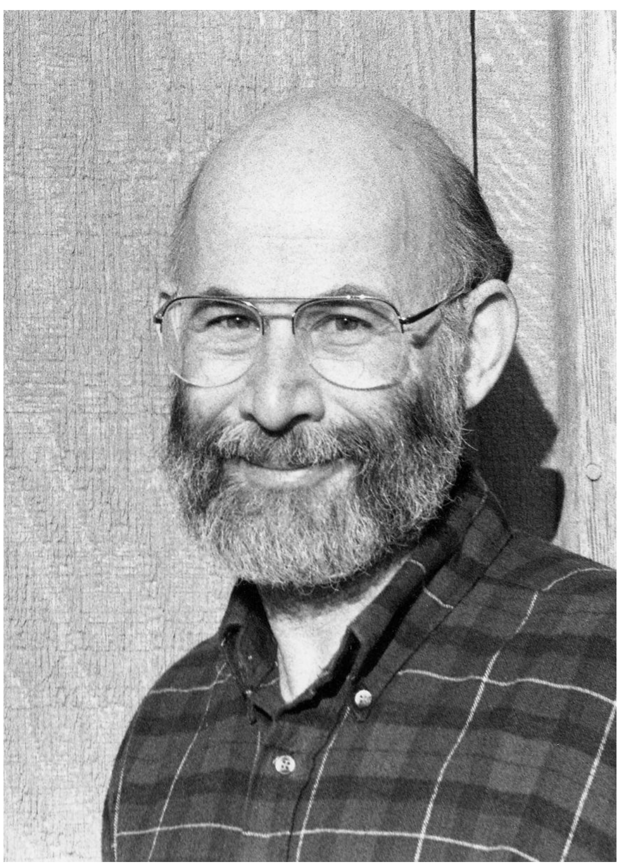

FIG. 4.15 Dr. Allen Neuringer.
Source: Photograph by Vera Jagendorf, Portland, OR. Published with permission.

At this point in the experiment, the researcher introduced a critical comparison phase where the birds were reinforced at the end of some of the 8-peck trials, but the presentation of food did not depend on the bird's variability.

Reinforcements were now tied or *yoked* to the earlier pattern of reinforcements delivered in the VAR phase. If a bird had been reinforced on a given trial of the VAR session, then the equivalent trial in the yoked phase also ended with reinforcement—regardless of whether the lag contingency had been met. By yoking each bird's reinforcements in the two phases, the researchers ensured that the allocation of reinforcements was identical for a given bird, and that the only difference was that variable sequences of some lag were *required* for reinforcement in the VAR phase, but not in the yoked phase, where variability was simply *permitted*. Would response variability be higher when variability produced reinforcement than when reinforcement occurred independently of variability? The answer was a clear yes. Birds were highly variable when reinforced for varying (VAR), but tended to repeat pecks to one of the keys when reinforcement did not require response variability (yoked). One conclusion is that variability is an operant dimension of behavior (much like force of response or speed of response) regulated by the contingency of reinforcement.

Reinforcement and Problem Solving: An Assessment

The current evidence indicates that *variability is an operant* that increases with reinforcement of behavioral variation (Lee, Sturmey, & Fields, 2007; Neuringer, 2002, 2009). To date, the reinforcement of variability has been shown in a number of species, including pigeons, dolphins, rats, and human adults and children (Goetz & Baer, 1973; van Hest, van Haaren, & van de Poll, 1989; Machado, 1997; Neuringer, 1986; Pryor, Haag, & O'Reilly, 1969; Stokes, Mechner, & Balsam, 1999). In addition, different experimental procedures have been used to produce variability with a number of different response forms (Blough, 1966; Eisenberger & Shanock, 2003; Goetz & Baer, 1973; Machado, 1989; Morgan & Neuringer, 1990; Odum, Ward, Barnes, & Burke, 2006; Pryor et al., 1969). Variability, constraints by task and contingencies, and artistic creativity have also been of interest (Stokes, 2001). More recently, an experiment with pigeons manipulated reinforcement magnitude for 4-peck sequences to left and right keys in a complex schedule arrangement (multiple schedule). The results showed that large reinforcers disrupted the reinforcement of variability, inducing a high level of behavioral repetition as the time to reinforcement approached (Doughty, Giorno, & Miller, 2013).

In summary, Barry Schwartz argued that reinforcement produces behavioral inflexibility and rigidity. In contrast, the research of Allen Neuringer and other behavior analysts suggested that response stereotypy is not an inevitable outcome of reinforcement. If the contingencies of reinforcement support stereotyped behavior, then this will occur. On the other hand, contingencies may generate novel, even creative, sequences of behavior if these patterns result in reinforcement (Neuringer, 2004; Neuringer & Jensen, 2013; Machado, 1989, 1992, 1997). Generally, a close analysis of the contingencies is required in problem-solving situations because "what you reinforce is what you get" (stereotypy or variability).

EXTINCTION

The procedure of *withholding reinforcement* for a previously reinforced response is called **extinction**. Skinner (1938) conducted the first extensive study of extinction and its related principles. To produce extinction, you would disconnect the food hopper after the bird had been reinforced for key